International Journal of Master of Engineering Research and Technology, 2394-6172-(O)
Volume-5, Issue-9, July 2018

A REVIEW OF USER BEHAVIOR BASED SECURITY ANALYSIS IN CLOUD COMPUTING

Aadesh Dangi
Research Scholar M. Tech.
Computer Science & Engineering
Surabhi College of Engineering & Technology
Bhopal M.P. India
Dangiaadesh2@gmail.com

Mr. Rakesh Kumar Lodhi
Assistant Professor
Computer Science & Engineering
Surabhi College of Engineering & Technology
Bhopal M.P. India
Rakeshlodhi21@gmail.com

ABSTRACT

Security is essential pillars of public cloud computing environments. In the public cloud environment provide some facility for access of data over the cloud network. The lack of standard protocol and business model the cloud environment always faced a security threat. For the analysis of security threats, the user behavior process is useful techniques. The user behavior process predicts the user process work over the cloud network. For the analysis of user behaviors used different algorithms such as association rule mining and many more data mining techniques. In this paper presents the review of user behavior security analysis based on different data mining techniques.

Keywords: - Cloud Computing, Data Security, User Behavior Data Mining.

INTRODUCTION

The advantages of cloud computing have been recognized by the information technology industry, but the extensive using of cloud computing is facing great challenges. to store confidential information and important business data into the cloud computing platform requires a lot of courage, as the user aware the security risk of the current cloud computing environment. the initial design of cloud computing within a narrow range, the resource a cloud computing resource sharing is under a closed environment, the basic security protections like firewall and traditional a cloud computing access control technology can satisfy the security requirement. but, with the development and expansion of cloud computing, the cloud resources are more and more exposure on the internet, the number and the scale of attacks are exponential increasing [1]. the security requirement was upgraded from traditional closed security a cloud computing accesses control into a dynamic, open security a cloud computing access control. security problems of cloud computing in data security and user privacy protection directly related to the further development of cloud computing. Many researchers and practitioners work on identifying cloud threats, vulnerabilities, attacks, and other security and privacy issues, in addition to providing countermeasures in the form of frameworks, strategies, recommendations, and service-oriented architectures. Additionally, efforts in other domains such as ad hoc networks have been tuned to address the emerging security problems in the clouds. Many researchers have addressed single attributes of cloud computing security such as data integrity, authentication vulnerabilities, auditing, etc. provide surveys that cover specific areas of cloud security concerns and proposed solutions. In [3,7], the authors briefly and broadly discuss cloud security issues involving data, applications and virtualization. The authors in [12] discuss similar cloud security issues but with deeper investigations. In [3,5], the authors present surveys on cloud security requirements such as confidentiality, integrity, transparency, availability, accountability, and assurance. The analysis of data over cloud network facility the service of cloud auditing process. The cloud auditing process generates the huge amount of data. these data process for the behavior analysis. In behavior analysis used data mining algorithm for the process of predication. Section II discusses about user behavior of cloud computing overview and description, Section III discusses about the related work. Section IV discusses problem formulation and finally, concluded in section V.

II USER BEHAVIOR IN CLOUD COMPUTING

User behavior analysis is new area of research in cloud computing environments. The user behavior analysis

generates the security threats risk level and market demand of cloud-based services. some points focus the process of cloud user behaviors[5].

- The number of users is much greater. Network technology is mainly to solve the problem of sharing information resources in different organizations. Cloud computing is not only able to achieve the sharing of resources but also able to achieve the integration of large-scale scalable computing, storage, data, and application of distributed computing for collaborative work of super computing model. In addition, it is bound to accommodate and attract more users[6].
- User behavior is more diverse. Cloud computing environment can provide more services. The users will have more interactive behavior in the use of these services. Therefore, the access control of interactive behavior is more important[6].
- Security issues are more prominent. Cloud computing faces a greater variety of security threats, such as availability of service, data lock-in, data confidentiality and audibility, data transfer bottlenecks, performance unpredictability, bugs in large-scale distributed systems, reputation fate sharing, privacy and security, and access control guarantee the user information, data, and the privacy and security of services is a problem that must be solved[8].

III RELATED WORK

In this section discuss the related work in the field of cloud data security and auditing of user behavior for the risk analysis of security measurements. Some work is discussing here. Chunye Zhao, Shanshan Tu, Haoyu Chen and Yongfeng Huang Et al. [1] this paper discussed user behavior-based scheme, by analyzing log data from cloud servers. firstly, they present the decryption rules of UB for operating system logs. then they put forward an association rule mining algorithm based on the long sequence frequent pattern to extract the UB. at last the result of experiment proves that our solution can implement the track and forensic of data leakage for the cloud security auditing.

S. Parkinsona, V. Somarakia and R. Warda Et al. [2] In this paper a novel method of modelling file system permissions which can be used by association rule mining techniques to identify irregular permissions is presented. This results in the creation of object-centric model as a by-product. This technique is then implemented and tested on Microsoft's New Technology File System permissions (NTFS). The results demonstrate that the technique is able to correctly identify irregularities with an average accuracy rate of 91%, minimizing the reliance on expert knowledge.

Vangipuram Radhakrishna, Puligadda Veereswara Kumar and Vinjamuri Janaki Et al. [3] The idea is to find normal temporal system call patterns and use these patterns to identify abnormal temporal system call patterns. For finding normal system call patterns, they use the concept of temporal association patterns. The reference sequence is used to obtain temporal association system call patterns satisfying specified dissimilarity threshold. To find similar (normal) temporal system call patterns, they apply our novel method which performs only a single database scan, reducing unnecessary extra overhead incurred when multiple scans are performed thus achieving space and time efficiency.

Gunupudi RajeshKumar, N Mangathayaru and G Narsimha Et al. [4] they mainly discuss the approach for intrusion detection by designing a distance measure which is designed by taking into consideration the conventional Gaussian function and modified to suit the need for similarity function. A Framework for intrusion detection is also discussed as part of this research.

Saurav Mallik, Anirban Mukhopadhyay and Ujjwal Maulik Et al. [5] they discussed a weighted rule-mining technique (say, RANWAR or rank-based weighted association rule-mining) to rank the rules using two novel rule-interestingness measures, viz., rank-based weighted condensed support (wcs) and weighted condensed confidence (wcc) measures to bypass the problem. These measures are basically depended on the rank of items (genes). Using the rank, they assign weight to each item. RANWAR generates much less number of frequent itemsets than the state-of-the-art association rule mining algorithms. The genes of the top rules are biologically validated by Gene Ontologies (GOs) and KEGG pathway analyses. Finally, the top rules evolved from RANWAR, that are not in Apriori, are reported.

Lichun Li, Rongxing Lu, Kim-Kwang Raymond Choo, Anwitaman Datta and Jun Shao Et al. [6] they focus on privacy-preserving mining on vertically partitioned databases. In such a scenario, data owners

wish to learn the association rules or frequent itemsets from a collective dataset and disclose as little information about their (sensitive) raw data as possible to other data owners and third parties. they then discussed a cloud-aided frequent itemset mining solution, which is used to build an association rule mining solution. Our solutions are designed for outsourced databases that allow multiple data owners to efficiently share their data securely without compromising on data privacy. Our solutions leak less information about the raw data than most existing solutions.

Meera Narvekar and Shafaque Fatma Syed Et al. [7] In this work, with the help of a new improved FP tree and a new Frequent Item set mining algorithm, they are able to save a lot of memory in terms of reducing the no. of conditional pattern bases and conditional FP trees generated. they are also able to reduce the no. of times the database is scanned. they conducted the experiments on the test dataset and found that the developed technique is efficient than the existing system. Also, the new improved FP tree gives the correlation among the items more clearly.

Xiuli Yuan Et al. [8] The improved Apriori algorithm achieves excellent performance by reducing the time consumed in transaction scanning for the generation of candidate itemsets and by reducing the number of transaction to be scanned. Generally, from view of time consumed, the T_Apriori performs much better as the group of transaction number increases and the value of minimum support increases. It reduced the time consumption by over 98% in average. Therefore, our improved Apriori algorithm is far more efficient than the original Apriori algorithm.

Sachin Kumar and Durga Toshniwal Et al. [9] they have used data mining techniques to analyze the data provided by EMRI in which they first cluster the accident data and further association rule mining technique is applied to identify circumstances in which an accident may occur for each cluster. The results can be utilized to put some accident prevention efforts in the areas identified for different categories of accidents to overcome the number of accidents.

Yongjun Ren, Jian Shen, Jin Wang, Jin Han and Sungyoung Lee Et al. [10] in this paper, they discussed an efficient mutual verifiable provable data possession scheme, which utilizes Diffie-Hellman shared key to construct the homomorphic authenticator. In particular, the verifier in our scheme is stateless and independent of the cloud storage service. It is worth noting that the presented scheme is very efficient compared with the previous PDP schemes, since the bilinear operation is not required.

Boyang Wang, Baochun Li and Hui Li Et al. [11] In this paper, they discussed a novel public auditing mechanism for the integrity of shared data with efficient user revocation in mind. By utilizing proxy re-signatures, they allow the cloud to re-sign blocks on behalf of existing users during user revocation, so that existing users do not need to download and re-sign blocks by themselves. In addition, a public verifier is always able to audit the integrity of shared data without retrieving the entire data from the cloud, even if some part of shared data has been re-signed by the cloud. Experimental results show that our mechanism can significantly improve the efficiency of user revocation.

Bob Duncan and Mark Whittington Et al. [12] they have identified fundamental weaknesses when undertaking cloud audit, namely the misconceptions surrounding the purpose of audit, what comprises a proper audit trail, what should be included, and how it should be achieved and maintained. A properly specified audit trail can provide a powerful tool in the armory against cyber-crime, yet it is all too easy to throw away the benefits offered by this simple tool through lack of understanding, incompetence, mis-configuration or sheer laziness. Of course, merely having an effective audit trail is not enough, they actually have to examine it regularly to realize the potential benefits it offers.

Mazhar Ali, Samee U. Khan and Athanasios V. Vasilakos Et al. [13] This survey details the security issues that arise due to the very nature of cloud computing. Moreover, the survey presents the recent solutions presented in the literature to counter the security issues. Furthermore, a brief view of security vulnerabilities in the mobile cloud computing are also highlighted. In the end, the discussion on the open issues and future research directions is also presented.

Chang Liu, Rajiv Ranjan, Chi Yang, Xuyun Zhang, Lizhe Wang and Jinjun Chen Et al. [14] they present a novel public auditing scheme named MuR-DPA. The new scheme incorporated a novel authenticated data structure based on the Merkle hash tree, which they name as MR-MHT. For support of full dynamic data updates, authentication of block indices and efficient verification of updates for multiple replicas at the same time, the level values of nodes in MR-MHT are generated in a top-down order, and all replica blocks for each data block are organized into a same replica sub-tree. Compared to existing integrity verification

and public auditing schemes, theoretical analysis and experimental results show that the MuR-DPA scheme can not only incur much less communication overhead for both update and verification of datasets with multiple replicas, but also provide enhanced security against dishonest cloud service providers.

IV DATA MINING

Data mining is play very important role in current growing rate of internet data. It is also a vital field of research in the field of pattern extraction and gathering of information on given database. The task of data mining is to extract useful knowledge for human users from a database. Here as the application of evolutionary computation to data mining is not always easy due to its heavy computation load especially in the case of a large database [15]. While data mining represents a significant advance in the type of analytical tools currently available, there are confines to its potential. One restriction is that although data mining can help reveal patterns and associations, it does not tell the user the value or significance of these patterns. These types of determinations must be made by the user. A second restriction is that while data mining can identify connections between behaviors and/or variables, it does not essentially identify a causal relationship. To be successful, data mining motionless requires skilled technical and analytical specialists who can structure the analysis and interpret the output that is created. Data mining is becoming more and more common in both the private and public sectors. Industries such as banking, insurance, medicine, and retailing commonly use data mining to decrease costs, improve research, and increase sales. Association rule mining is to find out association rules that satisfy the predefined minimum support and confidence from a given database. The problem is usually decomposed into two sub problems[9-11]. One is to find those item sets whose occurrences exceed a predefined threshold in the database; those item sets are called frequent or large item sets. The second problem is to generate association rules from those large item sets with the constraints of minimal confidence [15]. Suppose one of the large item sets is Lk, Lk = {I1, I2… Ik}, association rules with this item sets are generated in the following way: the first rule is {I1, I2… Ik-1} ⇒ {Ik}, by checking the confidence this rule can be determined as interesting or not. Then other rule are generated by Deleting the last items in the antecedent and inserting it to the consequent, further the confidences of the new rules are checked to determine the interestingness of them[8]. Those processes iterated until the antecedent becomes empty. Since the second sub problem is quite straight forward, most of the researches focus on the first sub problem. The first sub-problem can be further divided into two sub-problems: candidate large item sets generation process and frequent item sets generation process. We call those item sets whose support exceed the support threshold as large or frequent item-sets, those item sets that are expected or have the hope to be large or frequent are called candidate item sets[12].

V CONCLUSION & FUTURE SCOPE

In this paper present the review of user behaviors-based security analysis of cloud computing. The user behaviors are new area of analysis of security threats over the cloud network. For the analysis of cloud data used data mining techniques. For the analysis of data used relation constraints function. For this process used association rule mining algorithm. the association rule mining algorithm is good algorithm for user behaviors analysis. we present and evaluate the effectiveness of the state-of-the-art general countermeasures for cloud security attacks including user behavior systems, autonomous systems, and federated identity management systems. We also highlight the shortcomings of these systems that include the high communication and computation overhead and the detection efficiency and coverage. In future improved the rule mining algorithm and enhanced the process of cloud security analysis.

REFERENCES

[1] Chunye Zhao, Shanshan Tu, Haoyu Chen and Yongfeng Huang "Efficient association Rule mining based on user behaviour for cloud security auditing", IEEE, 2016, Pp 145-149.

[2] S. Parkinsona, V. Somarakia and R. Warda “Auditing file system permissions using Association Rule Mining”, IEEE, 2016, Pp 1-17.

[3] Vangipuram Radhakrishna, Puligadda Veereswara Kumar and Vinjamuri Janaki “A Novel Similar Temporal System Call Pattern Mining for Efficient Intrusion Detection”, Journal of Universal Computer Science, 2016, Pp 475-493.

[4] Gunupudi RajeshKumar, N Mangathayaru and G Narsimha “Intrusion Detection – A Text Mining Based Approach”, IJCSIS, 2016, Pp 76-88.

[5] Saurav Mallik, Anirban Mukhopadhyay and Ujjwal Maulik "RANWAR: Rank-Based Weighted Association Rule Mining from Gene Expression and Methylation Data", IEEE, 2015, Pp 1-8.

[6] Lichun Li, Rongxing Lu, Kim-Kwang Raymond Choo, Anwitaman Datta and Jun Shao "Privacy-Preserving Outsourced Association Rule Mining on Vertically Partitioned Databases", IEEE, 2016, Pp 1-15.

[7] Meera Narvekar and Shafaque Fatma Syed "An optimized algorithm for association rule mining using FP tree", Procedia Computer Science, 2015, Pp 101 – 110.

[8] Xiuli Yuan "An Improved Apriori Algorithm for Mining Association Rules", Advances in Materials, Machinery, Electronics I, 2016, Pp 1-7.

[9] Sachin Kumar and Durga Toshniwal "Analysing Road Accident Data Using Association Rule Mining", IEEE, 2015, Pp 1-6.

[10] Yongjun Ren, Jian Shen, Jin Wang, Jin Han and Sungyoung Lee "Mutual Verifiable Provable Data Auditing in Public Cloud Storage", Journal of Internet Technology, 2016, Pp 317-323.

[11] Boyang Wang, Baochun Li and Hui Li "Public Auditing for Shared Data with Efficient User Revocation in the Cloud", IEEE, 2015, Pp 1-9.

[12] Bob Duncan and Mark Whittington "Enhancing Cloud Security and Privacy: The Power and the Weakness of the Audit Trail", CLOUD COMPUTING, 2016, Pp 125-130.

[13] Mazhar Ali, Samee U. Khan and Athanasios V. Vasilakos "Security in cloud computing: Opportunities and challenges", Information Sciences, 2015, Pp 357–383.

[14] Chang Liu, Rajiv Ranjan, Chi Yang, Xuyun Zhang, Lizhe Wang and Jinjun Chen "MuR-DPA: Top-down Levelled Multi-replica Merkle Hash Tree Based Secure Public Auditing for Dynamic Big Data Storage on Cloud", IEEE, 2016, Pp 1-12.

[15] By Chi-Ren Shyu1,2, Matt Klaric1,2, Grant Scott1,2, and Wannapa Kay Mahamaneerat1 Knowledge Discovery by Mining Association Rules and Temporal-Spatial Information from Large-Scale Geospatial Image Databases 0-7803-9510IEEE 2006.